



App. No. 09/833,017
Inventor: Cvitkovitch et al.
Atty. Dkt: 2224-01301 (formerly 1889-00401)
Replacement Sheet 1 of 19

Figure 1

Streptococcus mutans
***ComCDE* Operon**

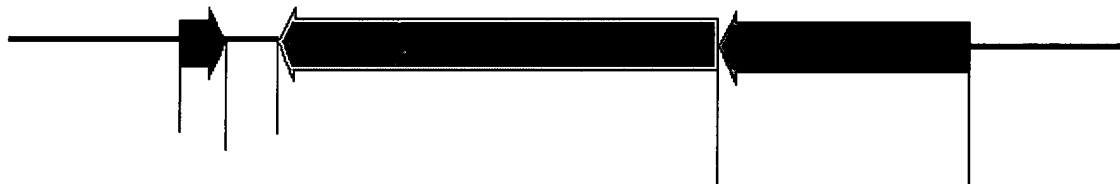


Figure 2A.

[ATGAAAAAACACTATCATTTAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAGA
GATTATCATTTGGCGGA (AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTT
TTACACAAGCTTTGGGAAAA)] TAA

Figure 2B.

AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTTTACACAAGCTTTGGGA
AAA

Figure 2C.

[ATGAATGAAGCCTTAATGATACTTTCAAATGGTTTATTAACCTTATCTAACCGT
TCTATTTCTCTTGTTTCTATTTTCTAAGGTAAGTAATGTCACCTTTATCGAAAAA
GGAATTAACCTTTTTTTCGATAAGCAATTTTCTGATAATGATTGCTGTTACGA
TGGTGAACGTAAACCTGTTTTATCCTGCAGAGCCTCTTTATTTTATAGCTTTAT
CAATTTATCTTAATAGACAGAATAGTCTTTCTCTAAATATATTTTATGGTCTGC
TGCCTGTTGCCAGTTCTGACTTGTTTAGGCGGGCAATCATATTCTTTATCTTGG
ATGGAACCTCAAGGAATTGTAATGGGCAGTAGCATTATAACACCTATATGAT
CGAGTTTGCAGGAATAGCGCTAAGTTACCTCTTTCTCAGTGTGTTCAATGTTG
ATATTGGTTCGACTTAAAGATAGTTTGACCAAGATGAAGGTCAAAAAACGCTT
GATTCCAATGAATATTACTATGCTTCTATACTACCTTTTAATACAGGTATTGT
ATGTTATAGAGAGTTATAATGTGATACCGACTTTAAAATTTTCGTAAATTTGTC
GTTATTGTCTATCTTATTTTATTTTGGATTCTGATCTCATTTTAAAGCCAATATA
CCAAACAAAAGGTTCAAAATGAGATAATGGCACAAAAGGAAGCTCAGATTC
GAAATATCACCCAGTATAGTCAGCAAATAGAATCTCTTTACAAGGATATTCG
AAGTTTCCGCCATGATTATCTGAATATTTTAACTAGCCTCAGATTAGGCATTG
AAAATAAAGATTTAGCTAGTATTGAAAAGATTTACCATCAAATCTTAGAAAA
AACAGGACATCAATTGCAGGATACCCGTTATAATATCGGCCATCTAGCTAAT
ATTCAAAACGATGCTGTCAAGGGTATCTTGTCAGCAAAAATCTTAGAAGCTC
AGAATAAAAAGATTGCTGTCAATGTAGAAGTCTCAAGTAAAATACAACTGCC
TGAGATGGAGTTGCTTGATTTCAATACCATACTTTCTATCTTGTTGATAATGC
CATTGAGGCTGCTTTCGAATCATTAATCCTGAAATTCAGTTAGCCTTTTTTA
AGAAAAATGGCAGTATAGTCTTTATCATTCAGAATTCACCAAAGAAAAACA
AATAGATGTGAGTAAAATTTTAAAGAAAATCTATTCCACTAAAGGCTCCAAT
CGCGGTATTGGTTTAGCAAAGGTGAATCATATTCTTGAACATTATCCCAAAAC
CAGTTTACAAACAAGCAATCATCATCATTTATTCAAGCAACTCCTAATAATAA
AA]TAG

Figure 2D.

[ATGATTTCTATTTTTGTATTGGAAGATGATTTTTTACAACAAGGACGTCTTGAAACCAC
CATTGCAGCTATCATGAAAGAAAAAATTGGTCTTATAAAGAATTGACTATTTTTGGAAA
ACCACAACAACCTTATTGACGCTATCCCTGAAAAGGGCAATCACCAGATTTTCTTTTTGGA
TATTGAAATCAAAAAAGAGGAAAAGAAAGGACTGGAAGTAGCCAATCAGATTAGACAGCA
TAATCCTAGTGCAGTTATTGTCTTTGTACGACACATTCTGAGTTTATGCCCCTCACTTT
TCAGTATCAGGTATCTGCTTTGGATTTTATTGATAAATCTTTGAATCCTGAGGAGTTCTC
CCACCGCATTGAATCAGCGCTGTATTATGCTATGGAAAACAGCCAGAAGAATGGTCAATC
AGAGGAACTTTTTATTTTCCATTCATCTGAAACTCAGTTTCAGGTCCCTTTTGCTGAGAT
TCTGTATTTTGAACATCTTCAACAGCCCATAAGCTCTGCCTTTATACTTATGATGAACG
GATTGAATTCTACGGCAGTATGACTGACATTGTTAAATGGATAAGAGACTTTTTCAGTG
CCATCGCTCTTTTATTGTCAATCCTGCCAATATTACCCGTATTGATCGGAAAAAACGCTT
GGCCTATTTTCGAAATAATAAGTCTTGTCTTATTTACGAACTAAGTTAACAAAACCTGAG
AGCTGTGATTGCTGATCAAAGGAGAGCAAAA] TGA

Figure 3

A.

MKKTLSLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK

B.

MNEALMILSNGLLTyLTVLFLFLFSKVSNTLSKKELTLFSISNFLIMIAVTMVNVNL
FYPAEPLYFIALSIYLNQRNSLSLNI FYGLLPVASSDLFRRAI IFFILDGTQGIVMGSS
IITTYMIEFAGIALSYLFLSVFNVDIGRLKDSLTKMKVKKRLIPMNITMLLYYLLIQVL
YVIESYNVIPTLKFRKFPVVIVYLILFLILISFLSQYTKQKVQNEIMAQKEAQIRNITQY
SQQIESLYKDIRSFRHDYLNILTSRLGIENKDLASIEKIYHQILEKTGHQLQDTRYNI
GHLANIQNDAVKGILSAKILEAQNKIAVNVEVSSKIQLPEMELLDFFITILSILCDNAI
EAAFESLNPEIQLAPFKKNGSIVFI IQNSTKEKQIDVSKI FKENYSTKGSNRGIGLAKV
NHILEHYPKTSLQTSNHHHLFKQLLI K

C.

MISIFVLEDDFLQQGRLETTIAAIMKEKNWSYKELTIFGKPQQLIDAIPEKGNHQIFFL
DIEIKKEEKKGLEVANQIRQHNPSAVIVFVTTHSEFMPLTFQYQVSALDFIDKSLNPEE
FSHRIESALYYAMENSQKNGQSEELFIHSSSETQFQVPPAEILYFETSSTAHLCLYTY
DERIEFYGSMTDIVKMDKRLFQCHRSFIVNPANITRIDRKKRLAYFRNNKSLISRTKL
TKLRAVIADQRRAK

Figure 4

A.

```
BM71 CSP      1 MKKTPSLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
GB14 CSP      1 MKKTLCLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
H7 CSP        1 MKKTLCLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
JH1005 CSP    1 MKKTLCLKNDKFKEIKTDELEIIIGSGTLSTFFRLFNRSFTQA   43
LT11 CSP      1 MKKTLCLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
NG8 CSP       1 MKKTLCLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
UAB159 CSP    1 MKKTLCLKNDKFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
*****
```

B.

```
consensus: 1 MKKTLCLKNDKFKEIKTDELEIIIGG SGSLSTFFRLFNRSFTQALGK 46
predicted cleavage site:
```

Figure 5

SGSLSTFFRLFNRSFTQALGK

Figure 6

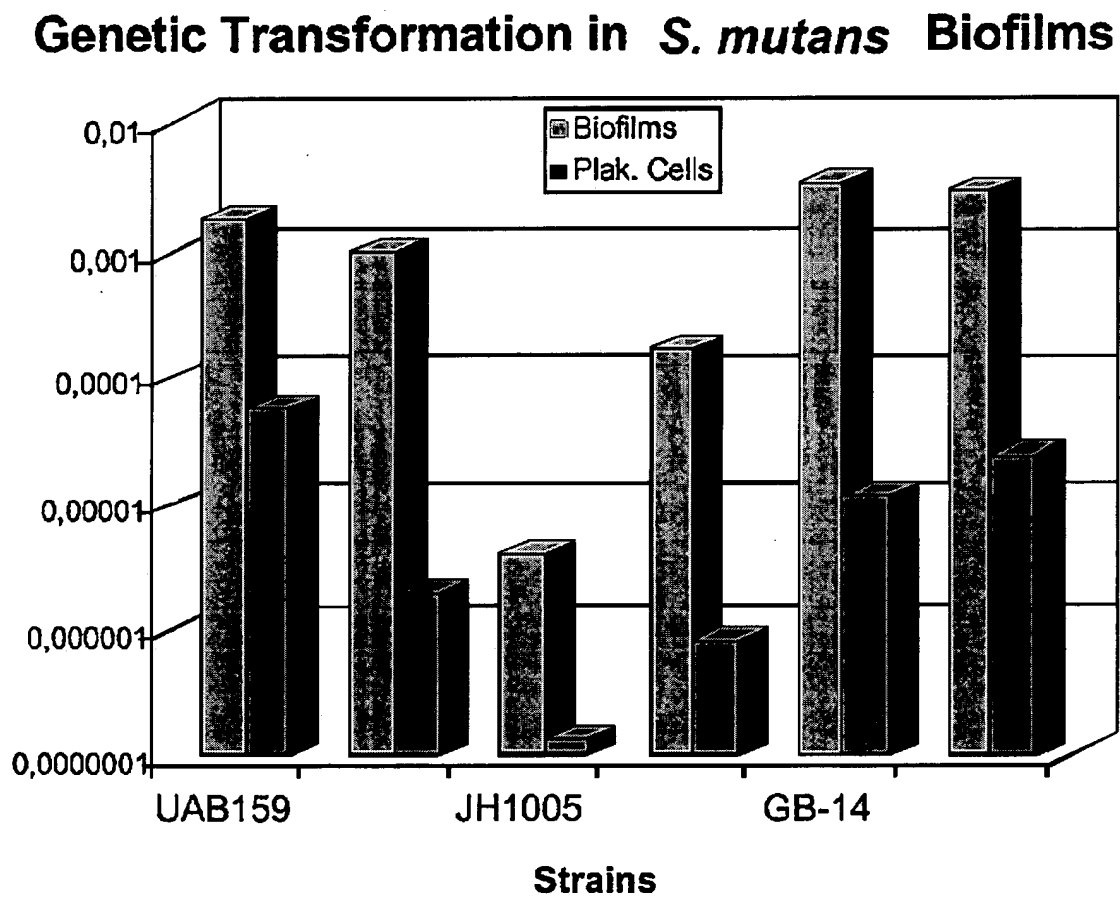


Figure 7

Strain	Peptide added Number of Transformants/Recipients	No peptide Number of Transformants/Recipients
UAB15	4.65×10^{-1}	1.78×10^{-6}
JH1005 ²	6.98×10^{-2}	0

¹The final concentration of SCSF used was 500 ng/ml.

The strain contains a nonsense mutation in the *comC* gene encoding the CSP.

Figure 8

ComC region

ComC Primer Pair: F5-B5

[F5] 23406-23424 5'- AGTTTTTTTGTCTGGCTGCG -3'

19 nt forward primer

pct G+C: 47.4 Tm: 50.5

[B5] 24056-24037 5'- TCCACTAAAGGCTCCAATCG -3'

20 nt backward primer

pct G+C: 50.0 Tm: 51.9

651 nt product for F5-B5 pair (23406-24056)

Optimal annealing temp: 50.3

pct G+C: 30.9 Tm: 71.5

ComD region

ComD Primer Pair: F1-B1

[F1] 392-415 5'- CGCTAAGTTACCTCTTTCTCAGTG -3'

24 nt forward primer

pct G+C: 45.8 Tm: 51.6

[B1] 683-663 5'- GCTTCCTTTTGTGCCATTATC -3'

21 nt backward primer

pct G+C: 42.9 Tm: 50.8

292 nt product for F1-B1 pair (392-683)

Optimal annealing temp: 49.5

pct G+C: 30.8 Tm: 70.2

ComE region

ComE Primer Pair: F1-B1

[F1] 145-165 5'- CCTGAAAAGGGCAATCACCAG -3'

21 nt forward primer

pct G+C: 52.4 Tm: 55.9

[B1] 606-585 5'- GCGATGGCACTGAAAAAGTCTC -3'

22 nt backward primer

pct G+C: 50.0 Tm: 55.4

462 nt product for F1-B1 pair (145-606)

Optimal annealing temp: 53.6

pct G+C: 38.3 Tm: 74.1

Figure 9A

Sequence Range: 1 to 2557

```

      10      20      30      40      50
ACATTATGTGTCTTAAGGAAAATATTACTTTTTCAAGAAAATCCATGATT
TGTAATACACAGGATTCCTTTTATAATGAAAAGTTCTTTTAGGTACTAA
      <K K L F I W S K
      <
      60      70      80      90     100
TTTTCATAAAAATAGTATACTAATTATAATCAAAAAAGGAGATATAAA
AAAAGTATTTTTATCATATGATTAATATTAGTTTTTCTCTATATTT
      <K M F F L I S I I I L F L L Y L
      <
      110     120     130     140     150
ATGAAAAAACACTATCATTAAAAATGACTTTAAAGAAATTAAGACTGA
TACTTTTTTTGTGATAGTAATTTTTTACTGAAATTTCTTTAATTCTGACT
      M K K T L S L K N D F K E I K T D>
      ORF RF [2] >
<I F F V S D N F F S K L S I L V S
      <
      160     170     180     190     200
TGAATTAGAGATTATCATTGGCGGAAGCGGAAGCCTATCAACATTTTTCC
ACTTAATCTCTAATAGTAACCGCCTTCGCCTTCGGATAGTTGTAAAAAGG
      E L E I I I G G S G S L S T F F>
      ORF RF [2] >
<S N S I I M
      <
      210     220     230     240     250
GGCTGTTTAACAGAAGTTTTACACAAGCTTTGGGAAAATAAGATAGGCTA
CCGACAAATTGTCTTCAAATGTGTTCGAAACCCTTTTATTCTATCCGAT
      R L F N R S F T Q A L G K>
      ORF RF [2] >
      260     270     280     290     300
ACATTGGAATAAAACAAGGCTGGATTTATTATTCCAGCCTTTTTAAATGT
TGTAACCTTATTTTGTTCGGACCTAAATAATAAGGTGCGAAAAATTTACA
      310     320     330     340     350
AAAAATAAAAATACAGGGTTAAATAATCAAGTGTGCTGTCGTGGATGAGAA
TTTTATTTTATGTCCCAATTTATTAGTTCACACGACAGCACCTACTCTT
      360     370     380     390     400
GATAAACTATCTCTTAGAGAATAGGCCTCCTCTATTTTATTATTAGGAG
CTATTTTGATAGAGAATCTCTTATCCGGAGGAGATAAAATAATAATCCTC
      <K I I L L
      < ORF RF [
      410     420     430     440     450
TTGCTTGAATAAATGATGATGATTGCTTGTGTGTAAGTGGTTTTGGGAT
AACGAACCTATTTACTACTACTAACGAACAAACATTTGACCAAAACCCTA
      <Q K F L H H H N S T Q L S T K P Y

```

Figure 9B

```

< _____ ORF RF[4] C _____

      460      470      480      490      500
AATGTTCAAGAATATGATTCACCTTTGCTAAACCAATACCGCGATTGGAG
TTACAAGTTCTTATACTAAGTGGAAACGATT'TGGTTATGGCGCTAACCTC
<H E L I H N V K A L G I G R N S
< _____ ORF RF[4] C _____

      510      520      530      540      550
CCTTTAGTGGAATAGTTTCTTTAAAAATTTTACTCACATCTATTTGTTT
GGAAATCACCTTATCAAAAGAAAATTTTAAAAATGAGTGTAGATAAACAAA
<G K T S Y N E K F I K S V D I Q K
< _____ ORF RF[4] C _____

      560      570      580      590      600
TTCTTTGGTGGAAATCTGAATGATAAAGACTATACTGCCATTTTCTTAA
AAGAAACCACCTTAAGACTTACTATTTCTGATATGACGGTAAAAAGAATT
<E K T S N Q I I F V I S G N K K F
< _____ ORF RF[4] C _____

      610      620      630      640      650
AAAAGGCTAACTGAATTTTCAGGATTTAATGATTTCGAAAGCAGCCTCAATG
TTTTCCGATTGACTTAAAGTCCTAAATTACTAAGCTTTCGTCGGAGTTAC
M>
< _____ ORF RF[4] C _____
<F A L Q I E P N L S E F A A E I>

      660      670      680      690      700
GCATTATCACACAAGATAGAAAGTATGGTAATGAAATCAAGCAACTCCAT
CGTAATAGTGTGTTCTATCTTTCATACCATTACTTTAGTTTCGTTGAGGTA
A L S H K I E S M V M K S S N S I>
< _____ ORF RF[3] _____ >
<A N D C L I S L I T I F D L L E M>
< _____ ORF RF[4] C _____

      710      720      730      740      750
CTCAGGCAGTTGTATTTTACTTGAGACTTCTACATTGACAGCAATCTTTT
GAGTCCGTCAACATAAAATGAACTCTGAAGATGTAAGTGTGTTAGAAAA
S G S C I L L E T S T L T A I F>
< _____ ORF RF[3] _____ >
<E P L Q I K S S V E V N V A I K K>
< _____ ORF RF[4] C _____

      760      770      780      790      800
TATTCTGAGCTTCTAAGATTTTGTGCTGACAAGATACCCTTGACAGCATCG
ATAAGACTCGAAGATTCTAAAAACGACTGTTCTATGGGAAGTGTCTGTAGC
L F>
< _____ ORF RF[4] C _____
<N Q A E L I K A S L I G K V A D>

      810      820      830      840      850
TTTTGAATATTAGCTAGATGGCCGATATTATAACGGGTATCCTGCAATTG
AAAACCTTATAATCGATCTACCGGCTATAATATTGCCCATAGGACGTTAAC
<N Q I N A L H G I N Y R T D Q L Q>
< _____ ORF RF[4] C _____

```

Figure 9C

```

      860      870      880      890      900
ATGTCCTGTTTTTTTCTAAGATTTGATGGTAAATCTTTTCAATACTAGCTA
TACAGGACAAAAAAGATTCTAAACTACCATTTAGAAAAGTTATGATCGAT
<H G T K E L I Q H Y I K E I S A L
<_____ORF RF[4] C_____

      910      920      930      940      950
AATCTTTATTTTCAATGCCTAATCTGAGGCTAGTTAAAATATTCAGATAA
TTAGAAATAAAAGTTACGGATTAGACTCCGATCAATTTTATAAGTCTATT
<D K N E I G L R L S T L I N L Y
<_____ORF RF[4] C_____

      960      970      980      990     1000
TCATGGCGGAAACTTCGAATATCCTTGTAAGAGATTCTATTTGCTGACT
AGTACCGCCTTTGAAGCTTATAGGAACATTTCTCTAAGATAAACGACTGA
  M A E T S N I L V K R F Y L L T>
<D H R F S R I D K Y L S E I Q Q S
<_____ORF RF[4] C_____

     1010     1020     1030     1040     1050
ATACTGGGTGATATTTTGAATCTGAGCTTCCTTTTGTGCCATTATCTCAT
TATGACCCACTATAAAGCTTAGACTCGAAGGAAAACACGGTAATAGAGTA
  I L G D I S N L S F L L C H Y L I>
<Y Q T I N R I Q A E K Q A M I E N
<_____ORF RF[4] C_____

     1060     1070     1080     1090     1100
TTTGAACCTTTTGTGTTGGTATATTGGCTTAAAAATGAGATCAGAATCAAA
AAACTTGGAACAAACCATATAACCGAATTTTACTCTAGTCTTAGTTT
  L N L L F G I L A>
<Q V K Q K T Y Q S L F S I L I L
<_____ORF RF[4] C_____

     1110     1120     1130     1140     1150
AATAAAATAAGATAGACAATAACGACAAATTTACGAAATTTTAAAGTCGG
TTATTTTATCTATCTGTTATTGCTGTTTAAATGCTTTAAATTTTACAGCC
<F L I L Y V I V V F K R F K L T P
<_____ORF RF[4] C_____

     1160     1170     1180     1190     1200
TATCACATTATAACTCTCTATAACATAACAATACCTGTATTAAAAGGTAGT
ATAGTGTAATATTGAGAGATATTGTATGTTATGGACATAATTTTCCATCA
<I V N Y S E I V Y L V Q I L L Y Y
<_____ORF RF[4] C_____

     1210     1220     1230     1240     1250
ATAGAAGCATAGTAATATTCATTGGAATCAAGCGTTTTTTGACCTTCATC
TATCTTCGTATCATTATAAGTAACCTTAGTTTCGCAAAAACTGGAAGTAG
<L L M T I N M P I L R K K V K M
<_____ORF RF[4] C_____

     1260     1270     1280     1290     1300
TTGGTCAAACATATCTTTAAGTCGACCAATATCAACATTGAACACACTGAG
```

Figure 9D

```
AACCAGTTTGATAGAAATTCAGCTGGTTATAGTTGTAAC TTGTGTGACTC
<K T L S D K L R G I D V N F V S L
< _____ ORF RF[4] C _____

      1310      1320      1330      1340      1350
AAAGAGGTAAC TTAGCGCTATTCCTGCAAAC TCGATCATATAGGTGGTTA
TTTCTCCATTGAATCGCGATAAGGACGTTTGAGCTAGTATATCCACCAAT
<F L Y S L A I G A F E I M Y T T I
< _____ ORF RF[4] C _____

      1360      1370      1380      1390      1400
TAATGCTACTGCCCATTAACAATTCCTTGAGTTCCATCCAAGATAAAGAAT
ATTACGATGACGGGTAATGTTAAGGAAC TCAAGGTAGGTTCTATTTCTTA
<I S S G M V I G Q T G D L I F F
< _____ ORF RF[4] C _____
      <L E K L E M W S L S Y
      < _____

      1410      1420      1430      1440      1450
ATGATTGCCCGCCTAAACAAGTCAGAACTGGCAACAGGCAGCAGACCATA
TACTAACGGGCGGATTTGTTTCAGTCTTGACCGTTGTCCGTCGTCTGGTAT
<I I A R R F L D S S A V P L L G Y
< _____ ORF RF[4] C _____
      <S Q G G L C T L V P L L C C V M
      < _____

      1460      1470      1480      1490      1500
AAATATATTTAGAGAAAGACTATTCTGTCTATTAAGATAAATTGATAAAG
TTTATATAAATCTCTTTCTGATAAGACAGATAATTCTATTTAACTATTTT
<F I N L S L S N Q R N L Y I S L A
< _____ ORF RF[4] C _____

      1510      1520      1530      1540      1550
CTATAAAATAAAGAGGCTCTGCAGGATAAAACAGGTTTACGTTCCACCATC
GATATTTTATTTCTCCGAGACGTCCTATTTTGTCCAAATGCAAGTGGTAG
<I F Y L P E A P Y F L N V N V M
< _____ ORF RF[4] C _____

      1560      1570      1580      1590      1600
GTAACAGCAATCATTATCAGAAAAATTGCTTATCGAAAAAAGAGTTAATTC
CATTGTCGTTAGTAATAGTCTTTTAAACGAATAGCTTTTTTCTCAATTAAG
<T V A I M I L F N S I S F L T L E
< _____ ORF RF[4] C _____

      1610      1620      1630      1640      1650
CTTTTTCGATAAAGTGACATTACTTACCTTAGAAAAATAGAAACAAGAGAA
GAAAAAGCTATTTCACTGTAATGAATGGAATCTTTTATCTTTGTTCTCTT
<K K S L T V N S V K S F L F L L F
< _____ ORF RF[4] C _____

      1660      1670      1680      1690      1700
ATAGAACGGTTAGATAAGTTAATAAACCATTTGAAAGTATCATTAAGGCT
TATCTTGCCAATCTATTCAATTATTTGGTAAACTTTTCATAGTAATTCCGA
<L V T L Y T L L G N S L I M L A
< _____ ORF RF[4] C _____

      1710      1720      1730      1740      1750
```

Figure 9E

```

TCATTCATTTTGCCTCTCCTTTGATCAGCAATCACAGCTCTCAGTTTGTGTT
AGTAAGTAAAACGAGAGGAACTAGTCGTTAGTGTGAGAGTCAAAACAA
<E N M
<
<K A R R Q D A I V A R L K T
< ORF RF[5] C
1760 1770 1780 1790 1800
AACTTAGTTCGTGAAATAAGACAAGACTTATTATTTTCGAAAATAGGCCAA
TTGAATCAAGCACTTTATTCTGTTCTGAATAATAAAGCTTTTATCCGGTT
<L K T R S I L C S K N N R F Y A L
< ORF RF[5] C
1810 1820 1830 1840 1850
GCGTTTTTTCCGATCAATACGGGTAATATTGGCAGGATTGACAATAAAAG
CGCAAAAAGGCTAGTTATGCCATTATAACCGTCCTAACTGTTATTTTC
<R K K R D I R T I N A P N V I F S
< ORF RF[5] C
1860 1870 1880 1890 1900
AGCGATGGCACTGAAAAAGTCTCTTATCCATTTTAACAATGTCAGTCATA
TCGCTACCGTGACTTTTTTCAGAGAATAGGTAAAATTGTTACAGTCAGTAT
M A L K K S L I H F N N V S H>
ORF RF[1]
<R H C Q F L R K D M K V I D T M
< ORF RF[5] C
<V
<
1910 1920 1930 1940 1950
CTGCCGTAGAAATCAATCCGTTTCATCATAAGTATAAAGGCAGAGCTTATG
GACGGCATCTTAAGTTAGGCAAGTAGTATTTCATATTTCCGTCTCGAATAC
T A V E F N P F I I S I K A E L M>
ORF RF[1]
<S G Y F E I R E D Y T Y L C L K H
< ORF RF[5] C
<A T S N L G N M M L I F A S S I
< ORF RF[6] C
1960 1970 1980 1990 2000
GGCTGTTGAAGATGTTTCAAAATACAGAATCTCAGCAAAAGGGACCTGAA
CCGACAACCTTCTACAAAGTTTTATGTCTTAGAGTCGTTTTCCCTGGACTT
G C>
>
<A T S S T E F Y L I E A F P V Q F
< ORF RF[5] C
<P Q Q L H K L I C F R L L L S R F
< ORF RF[6] C
2010 2020 2030 2040 2050
ACTGAGTTTCAGATGAATGGAAAATAAAAAGTTCCTCTGATTGACCATTCT
TGACTCAAAGTCTACTTACCTTTTATTTTCAAGGAGACTAACTGGTAAG
<Q T E S S H F I F L E E S Q G N
< ORF RF[5] C
<S L K L H I S F L F N R Q N V M R
< ORF RF[6] C

```

Figure 9F

```

2060      2070      2080      2090      2100
TTCTGGCTGTTTTCCATAGCATAATACAGCGCTGATTCAATGCGGTGGGA
AAGACCGACAAAAGGTATCGTATTATGTCGCGACTAAGTTACGCCACCCT
<K Q S N E M A Y Y L A S E I R H S
<_____ORF RF[5] C_____
      <R A T K W L M
<_____ORF RF[6] C_____

2110      2120      2130      2140      2150
GAACTCCTCAGGATTCAAAGATTTATCAATAAAATCCAAAGCAGATACCT
CTTGAGGAGTCCTAAGTTTCTAAATAGTTATTTTAGGTTTCGTCTATGGA
<F E E P N L S K D I F D L A S V Q
<_____ORF RF[5] C_____

2160      2170      2180      2190      2200
GATACTGAAAAGTGAGGGGCATAAACTCAGAATGTGTCGTGACAAAGACA
CTATGACTTTTCACTCCCCGTATTTGAGTCTTACACAGCACTGTTTCTGT
      M C R D K D>
<Y Q F T L P M F E S H T T V F V
<_____ORF RF[5] C_____

2210      2220      2230      2240      2250
ATAACTGCACTAGGATTATGCTGTCTAATCTGATTGGCTACTTCCAGTCC
TATTGACGTGATCCTAATACGACAGATTAGACTAACCGATGAAGGTCAGG
N N C T R I M L S N L I G Y F Q S>
<I V A S P N H Q R I Q N A V E L G
<_____ORF RF[5] C_____

2260      2270      2280      2290      2300
TTTCTTTTCCTCTTTTTTTGATTTCATATCCAAAAAGAAAATCTGGTGAT
AAAGAAAAGGAGAAAAAACTAAAGTTATAGGTTTTTCTTTTAGACCACTA
F L F L F F D F N I Q K E N L V I>
<K K E E K K I E I D L F F I Q H N
<_____ORF RF[5] C_____

2310      2320      2330      2340      2350
TGCCCTTTTCAGGGATAGCGTCAATAAGTTGTTGTGGTTTTTCCAAAAATA
ACGGGAAAAAGTCCCTATCGCAGTTATTCAACAACACCAAAAGGTTTTTAT
A L F R D S V N K L L W F S K N>
<G K E P I A D I L Q Q P K G F I
<_____ORF RF[5] C_____

2360      2370      2380      2390      2400
GTCAATTCTTTATAAGACCAATTTTTTTCTTTTCATGATAGCTGCAATGGT
CAGTTAAGAAATATTCTGGTTAAAAAAAGAAAGTACTATCGACGTTACCA
S Q F F I R P I F F F H D S C N G>
      M I A A M V>
<T L E K Y S W N K E K M I A A I T
<_____ORF RF[5] C_____

2410      2420      2430      2440      2450

```

Figure 9G

```
GGTTTCAAGACGTCCTTGTTGTAAAAATCATCTTCCAATACAAAAATAG
CCAAAGTTCTGCAGGAACAACATTTTTTAGTAGAAGGTTATGTTTTTATC
G F K T S L L>
_____>
V S R R P C C K K S S S N T K I>
<T E L R G Q Q L F D D E L V F I S
<_____ORF RF[5] C_____>
2460      2470      2480      2490      2500
AAATCATTATTTCTCCTTTAATCTTCTATTTAGGTTAGCTGATTAACACT
TTTAGTAATAAAGAGGAAATTAGAAGATAAATCCAATCGACTAATTGTGA
E I I I S P L I F Y L G>
_____>
<T M
<_____
2510      2520      2530      2540      2550
ATACACAGAAAAGGTATAAAACGATATCACTCAATAAAATCTACTAACTT
TATGTGTCCTTTCCATATTTTGCTATAGTGAGTTATTTTAGATGATTGAA
AATAACC
TTATTGG
```

Figure 10

A.

ATGGAAGAAGATTTTGAATTTGTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAG
CCGTTATTACTTTATTAAAAATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGA
TTTTGCACCAATTATTAAGGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTAT
ATCTATTTTAAGACACGTTTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAG
TCAGAAACGTCGTTTTAATAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTT
TGTC AAGTGGCGGTATGCAATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTGCA
TATAAACAAGGTCTGAGTACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGA
GCACTTTTTGGGAAGGCAAAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTA
AGGAAAAA

B.

MEEDFEIVFNKVKPIVWKL SRYFFIKMWTREDWQQEGMLILHQLLREHPELEEDDTKLY
IYFKTRFSNYIKDVL RQQESQKRRFN RMSYEEVGEIEHCLSSGGMQLDEYILFRDSL LA
YKQGLSTEQELFERLVAGEHFLGRQSM LKDLRKKLSDFKEK

C.

GTAAATAAACAGCCAGTTAAGATGGGACATTTATGTCCTGTTCTTAAAGTCTTTTTTCG
TTTTATAATAATTTTATTATAAAAGGAGGTCATCGTAATAGATGGAAGAAGATTTTGAA
ATTGTTTTTAATAAGGTTAAGCCAATTGTATGGAAATTAAGCCGTTATTACTTTATTAA
AATGTGGACTCGTGAAGATTGGCAACAAGAGGGAATGTTGATTTTGCACCAATTATTAA
GGGAACATCCAGAATTAGAAGAGGATGATACAAAATTGTATATCTATTTTAAGACACGT
TTTTCTAATTACATTAAAGATGTTTTGCGTCAGCAAGAAAGTCAGAAACGTCGTTTTAA
TAGAATGTCTTATGAAGAAGTCGGTGAGATTGAACACTGTTTGTCAAGTGGCGGTATGC
AATTGGATGAATATATTTTATTTCGTGATAGTTTGCTTG CATATAAACAAGGTCTGAGT
ACTGAAAAGCAAGAGCTGTTTGAGCGCTTGGTAGCAGGAGAGCACTTTTTGGGAAGGCA
AAGTATGCTGAAAGATTTACGTAAAAAATTAAGTGATTTTAAGGAAAAATAGTTAAAAA
GGGAAAGAATGGAACATGTGATTGTACCATTCTTTTTGGTTGAAAATTAAGAAAAGTTA
TTATAAATTATTGGTTTAACATGCCATATTA

Figure 11A.

ATGAAACAAGTTATTTATGTTGTTTTAATCGTCATAGCCGTTAACATTCTCTTAGAGATT
ATCAAAAGAGTAACAAAAAGGGGAGGGACAGTTTCGTCATCTAATCCTTTACCAGATGGG
CAGTCTAAGTTGTTTTGGCGCAGACATTATAAGCTAGTACCTCAGATTGATACCAGAGAC
TGTGGGCCCGCAGTGCTGGCATCTGTTGCAAAGCATTACGGATCTAATTACTCTATCGCT
TATCTGCGGGAACCTCTCAAAGACTAACAAGCAGGGAACAACAGCTCTTGGCATTGTTGAA
GCTGCTAAAAAGTTAGGCTTTGAAACACGCTCTATCAAGGCGGATATGACGCTTTTTGAT
TATAATGATTTGACCTATCCTTTTATCGTCCATGTGATTAAAGGAAAAACGTCTGCAGCAT
TATTATGTCGTCTATGGCAGCCAGAATAATCAGCTGATTATTGGAGATCCTGATCCTTCA
GTAAAGGTGACTAGGATGAGTAAGGAACGCTTTCAATCAGAGTGGACAGGCCTTGCAATT
TTCCTAGCTCCTCAGCCTAACTATAAGCCTCATAAAGGTGAAAAAATGGTTTGTCTAAT
TTCTTCCCGTTGATCTTTAAGCAGAAAGCTTTGATGACTTATATTATCATAGCTAGCTTG
ATTGTGACGCTCATTGATATTGTCGGATCATACTATCTCCAAGGAATATTGGACGAGTAC
ATTCTGATCAGCTGATTTCACTTTAGGAATGATTACGATTGGTCTGATAATAACCTAT
ATTATCCAGCAGGTCATGGCTTTTGCAAAAGAATACCTCTTGGCCGTA CT CAGTTTGCCT
TTAGTCATTGATGTTATCCTGTCTTATATCAAACATATTTTTACGCTTCCTATGTCTTTC
TTTGCGACAAGGCGAACAGGAGAAATCACGTCTCGTTTTACAGATGCCAATCAGATTATT
GATGCTGTAGCGTCAACCATCTTTTCAATCTTTTATAGATATGACTATGGTAATTTTGGTT
GGTGGGGTTTTGTTGGCGCAAAACAATAACCTTTTCTTCTAACCTTGCTCTCCATTCCG
ATTTATGCCATCATTATTTTTGCTTTCTTGAAACCCTTTGAGAAAATGAATCACGAAGTG
ATGAAAGCAATGCTGTGGTAAGTTCTTCTATCATTTGAAGATATCAATGGGATGGAAACC
ATTAAATCACTCACAAAGTGAGTCCGCTCGTTATCAAAACATTGATAGTGAATTTGTTGAT
TATTTGGAGAAAACTTTAAGCTACACAAGTATAGTGCCATTCAAACCGCATTA AAAAGC
GGTGCTAAGCTTATCCTCAATGTTGTCAATTCTCTGGTATGGCTCTCGTCTAGTTATGGAT
AATAAAATCTCAGTTGGTCAGCTTATCACCTTTAATGCTTTGCTGTCTTATTTCTCAAAT
CCAATTGAAAATATTATCAATCTGCAATCCAACTGCAGTCAGCTCGCGTTGCCAATACA
CGTCTTAATGAGGTCTATCTTGTCGAATCTGAATTTGAAAAAGACGGCGATTTATCAGAA
AATAGCTTTTTAGATGGTGATATTTCTGTTTGAAAATCTTTCTTATAAATATGGATTTGGG
CGAGATACCTTATCAGATATTAATTTATCAATCAAAAAAGGCTCCAAGGTCAGTCTAGTT
GGAGCCAGTGTTCTGGTAAACAACCTTTGGCTAAACTGATTGTCAATTTCTACGAGCCT
AACAAGGGGATTGTTCGAATCAATGGCAATGATTTAAAAGTTATTGATAAGACAGCTTTG
CGGCGGCATATTAGCTATTTGCCGCAACAGGCCTATGTTTTTAGTGGCTCTATTATGGAT
AATCTCGTTTTAGGAGCTAAAGAAGGAACGAGTCAGGAAGACATTATTCGTGCTTGTGAA
ATTGCTGAAATCCGCTCGGACATTGAACAAATGCCTCAGGGCTATCAGACAGAGTTATCA
GATGGTGCCGGTATTTCTGGCGGTCAAAACAGCGGATTGCTTTAGCTAGGGCCTTATTA
ACACAGGCACCGGTTTTGATTCTGGATGAAGCCACCAGCAGTCTTGATATTTTGACAGAA
AAGAAAATTATCAGCAATCTCTTACAGATGACGGAGAAAAACAATAATTTTTGTGCCCAC
CGCTTAAGCATTTTACAGCGTACTGACGAAGTCATTGTCTATGGATCAGGGAAAAATTGTT
GAACAAGGCACTCATAAGGAACCTTTAGCTAAGCAAGGTTTCTATTATAACCTGTTTAAT

Figure 11B.

MKQVIYVVLIVIAVNILLEIIKRVTKRGGTVSSSNPLPDGQSKLFWRRHYKLVPQIDTRD
CGPAVLASVAKHYGSNYSIAYLRELSKTNKQGTTALGIVEAAKKLGFETRSIKADMTLFD
YNDLTYPFIVHVIKGRQLQHYYVVYGSQNNQLIIGDPDPSVKVTRMSKERFQSEWTGLAI
FLAPQPNYKPHKGEKNGLSNFFPLIFKQKALMTYIIIASLIVTLIDIVGSYYLQGILDEY
IPDQLISTLGMITIGLIITYIIQQVMAFAKEYLLAVLSLRLVIDVILSYIKHIFTLPMSF
FATRRTGEITSRFTDANQIIDAVASTIFSIFLDMTMVILVGGVLLAQNNLFFLTLLSIP
IYAIIFAFKLPFEKMNHEVMESNAVVSSSIIEDINGMETIKSLTSESARYQNIDSEFVD
YLEKNFKLHKYSAIQTALKSGAKLILNVVILWYGSRLVMDNKISVGQLITFNALLSYFSN
PIENIINLQSKLQ SARVANTRLNEVYLVESEFEKDGDLSENSFLDGDISFENLSYKYGFG
RDTLSDINLSIKKGSKVS LVGASGSGKTTLAKLIVNFYEPNKGIVRINGNDLKVIDKTAL
RRHISYLPQQAYVFGSGSIMDNLVLGAKEGTSQEDIIRACEIAEIRSDIEQMPQGYQTELS
DGAGISGGQKQRIALARALLTQAPVILIDEATSSLDILTEKKIISNLLQMTEKTIIFVAH
RLSISQRTDEVIVMDQGKIVEQGTHKELLAKQGFYYNLFN

Figure 11C.

ATGGATCCTAAATTTTTACAAAGTGCAGAATTTTATAGGAGACGCTATCATAATTTTGC
ACACTATTAATTGTTCTTTGGTCTGCTTGATTATCTTCTTGGTCATATTCCTTTGTTTT
GCTAAAAAGAAATTACAGTGATTTCTACTGGTGAAGTTGCACCAACAAAGGTTGTAGAT
GTTATCCAATCTTACAGTGACAGTTCAATCATTAAAAATAATTTAGATAATAATGCAGCT
GTTGAGAAGGGAGACGTTTTAATTGAATATTCAGAAAATGCCAGTCCAAACCGTCAGACT
GAACAAAAGAATATTATAAAAGAAAGACAAAAACGAGAAGAGAAGGAAAAGAAAAAACAC
CAAAAGAGCAAGAAAAAGAAGAGTCTAAGAGCAAGAAAGCTTCCAAAGATAAGAAAAAG
AAATCGAAAGACAAGGAAAGCAGCTCTGACGATGAAAATGAGACAAAAAGGTTTCGATT
TTTGCTTCAGAAGATGGTATTATTCATACCAATCCCAAATATGATGGTGCCAATATTATT
CCGAAGCAAACCGAGATTGCTCAAACTATCCTGATATTCAAAAAACAAGAAAAGTGTTA
ATCACCTATTATGCTTCTTCTGATGATGTTGTTTCTATGAAAAGGGGCAAACCGCTCGT
CTTTCCTTGGAAGAAAAGGGAAATGACAAGGTTGTTATTGAAGGAAAATTAACAATGTC
GCTTCATCAGCAACTACTACTAAAAAGGAAATCTCTTTAAGGTACTGCCAAAGTAAAG
GTTTCTAAGAAAAATAGCAAACCTCATCAAGTATGGTATGACAGGCAAGACAGTCACTGTC
ATTGATAAAAAGACTTATTTTGATTATTTCAAAGATAAATTACTGCATAAAATGGATAAT

Figure 11D.

MDPKFLQSAEFYRRRYHNFATLLIVPLVCLIIIFLVIFLCFAKKEITVISTGEVAPTKVVD
VIQSYSDSSIKNNDNNAAVEKGDVLEISENASPNRQTEQKNIKERQKREEKEKKKH
QKSKKKKSKSKKASKDKKKKSKDKESSDDENETKKVSI FASEDGI IHTNPKYDGANII
PKQTEIAQIYPDIQKTRKVLITYYASSDDVSMKKGQTARLSLEKKGNDKVVEGKINNV
ASSATTTKKNLKFVTA KVVKVSKKNSKLIKGMTGKTVTVIDKKTYFDYFKDKLLHKMDN

Figure 12

